15-251

Great Theoretical Ideas in Computer Science

Polynomials, Lagrange, and Error-correction

Q: suppose P(x) has a rational root (a/b) over the rationals. Is $P(ab^{-1}) = 0$ over Z_p ?

(Let's assume both a, b in Z_p)

E.g., suppose $P(x) = 6x^2 - x - 1$
$= (2x - 1)(3x + 1)$.

Roots are 1/2, -1/3.

Roots over Z_{11} are $2^{-1} = 6$, $-3^{-1} = -4 = 7$

Q: P(x) has no root over the rationals. Does it have roots when working over Z_p?

Consider $P(x) = x^2 + 2x + 2$.

Over the reals, its roots are irrational.

Over Z_5, this is the same as $x^2 - 3x + 2$, which has roots 1,2 (both in Z_5)

The Single Most Important Theorem About Polynomials

A non-zero degree-d polynomial P(x) has at most d roots.

This fact has many applications…

Theorem:

Given pairs $(a_1, b_1), \ldots, (a_{d+1}, b_{d+1})$ of values there is at most one degree-d polynomial P(x) such that:
$P(a_k) = b_k$ for all k

P(1) = 0
P(3) = [illegible]
P(5) = 19

when we say "degree-d", we mean degree at most d.

we'll always assume $a_i \neq a_k$ for $i \neq k$

Theorem:

Given pairs $(a_1, b_1), \ldots, (a_{d+1}, b_{d+1})$ of values
there is at most one
degree-d polynomial P(x)
such that:
$P(a_k) = b_k$ for all k

do there exist d+1 pairs
for which there are
no such polynomials??

Revised Theorem:

Given pairs $(a_1, b_1), \ldots, (a_{d+1}, b_{d+1})$ of values
there is exactly one
degree-d polynomial P(x)
such that:
$P(a_k) = b_k$ for all k

The algorithm to construct P(x)
is called Lagrange Interpolation

Two different representations

$P(x) = c_d x^d + c_{d-1} x^{d-1} + \ldots + c_1 x^1 + c_0$
can be represented either by

a) its d+1 coefficients
$c_d, c_{d-1}, \ldots, c_2, c_1, c_0$

b) Its value at any d+1 points
$P(a_1), P(a_2), \ldots, P(a_d), P(a_{d+1})$
(e.g., P(1), P(2), …, P(d+1).)

Converting Between The Two Representations

Coefficients to Evaluation:

Evaluate P(x) at d+1 points

Evaluation to Coefficients:

Use Lagrange Interpolation

Now for some Lagrange Interpolation

Given pairs $(a_1, b_1), \ldots, (a_{d+1}, b_{d+1})$ of values
there is exactly one
degree-d polynomial P(x)
such that:
$P(a_k) = b_k$ for all k

Special case

What if the points were like:

$(a_1, 1)$
$(a_2, 0)$
$(a_3, 0)$
…
$(a_{d+1}, 0)$

Special case

Suppose we can get degree-d poly $h_1(x)$:
$h_1(a_1) = 1$
$h_1(a_t) = 0$ for all $t = 2,\ldots,d+1$

"switch" polynomial #1

Special case

Suppose we can get degree-d poly $h_1(x)$:
$h_1(a_1) = 1$
$h_1(a_t) = 0$ for all $t = 2,\ldots,d+1$

Then we can get degree-d poly $H_1(x)$:
$H_1(a_1) = b_1$
$H_1(a_t) = 0$ for all $t = 2,\ldots,d+1$

Just set $H_1(x) = b_1 * h_1(x)$

Special case

Suppose we can get degree-d poly $h_1(x)$:
$h_1(a_1) = 1$
$h_1(a_t) = 0$ for all $t = 2,\ldots,d+1$

Using same idea, get degree-d poly $H_k(x)$:
$H_k(a_k) = b_k$
$H_k(a_t) = 0$ for all $t \neq k$

Finally, $P(x) = \sum_k H_k(x)$

Hence, all we need to do

Given numbers $a_1, a_2, \ldots, a_{d+1}$

Build a degree-d "switch" poly $h_1(x)$:
$h_1(a_1) = 1$
$h_1(a_t) = 0$ for all $t = 2,\ldots,d+1$

construction by example

want a quadratic h with h(3) = 1, h(1)=0, h(6)=0

(say, in $Z_{11}[x]$)

Let's first get the roots in place:

$h(x) = (x-1)(x-6)$

Are we done? No! We wanted h(3) = 1

But h(3) = (3-1)(3-6) = -6

So let's fix that!

$h(x) = (-6)^{-1} (x-1)(x-6)$
$= 9 (x-1)(x-6)$

$9 * (-6) =_{11} 1$

done!

formally, the constructions

k-th “Switch” polynomial

$g_k(x) = (x-a_1)(x-a_2)\ldots(x-a_{k-1})(x-a_{k+1})\ldots(x-a_{d+1})$

Degree of $g_k(x)$ is: d

$g_k(x)$ has d roots: $a_1,\ldots,a_{k-1},a_{k+1},\ldots,a_{d+1}$

$g_k(a_k) = (a_k-a_1)(a_k-a_2)\ldots(a_k-a_{k-1})(a_k-a_{k+1})\ldots(a_k-a_{d+1})$

For all $i \neq k$, $g_k(a_i) = 0$

k-th “Switch” polynomial

$g_k(x) = (x-a_1)(x-a_2)\ldots(x-a_{k-1})(x-a_{k+1})\ldots(x-a_{d+1})$

$$h_k(x) = \frac{(x-a_1)(x-a_2)\ldots(x-a_{k-1})(x-a_{k+1})\ldots(x-a_{d+1})}{(a_k-a_1)(a_k-a_2)\ldots(a_k-a_{k-1})(a_k-a_{k+1})\ldots(a_k-a_{d+1})}$$

$h_k(a_k) = 1$

For all $i \neq k$, $h_k(a_i) = 0$

The Lagrange Polynomial

$$h_k(x) = \frac{(x-a_1)(x-a_2)\ldots(x-a_{k-1})(x-a_{k+1})\ldots(x-a_{d+1})}{(a_k-a_1)(a_k-a_2)\ldots(a_k-a_{k-1})(a_k-a_{k+1})\ldots(a_k-a_{d+1})}$$

$$P(x) = b_1\, h_1(x) + b_2\, h_2(x) + \ldots + b_{d+1}\, h_{d+1}(x)$$

P(x) is the <u>unique</u> polynomial of degree d such that $P(a_1) = b_1$, $P(a_2) = b_2$, …, $P(a_{d+1}) = b_{d+1}$

Example

Input: (5,1), (6,2), (7,9)

Want quadratic in $Z_{11}[x]$

Switch polynomials:

$h_1(x) = (x-6)(x-7)/(5-6)(5-7) = ½\,(x-6)(x-7)$

$h_2(x) = (x-5)(x-7)/(6-5)(6-7) = -\,(x-5)(x-7)$

$h_3(x) = (x-5)(x-6)/(7-5)(7-6) = ½\,(x-5)(x-6)$

$P(x) = 1 \times h_1(x) + 2 \times h_2(x) + 9 \times h_3(x)$

$= (3x^2 - 32x + 86)$

$= (3x^2 + x + 9)$ in $Z_{11}[x]$

the Chinese Remainder Theorem uses very similar ideas in its proof

Revised Theorem:

Given pairs (a_1, b_1), …, (a_{d+1}, b_{d+1}) of values there is <u>exactly one</u> degree-d polynomial P(x) such that:

$P(a_k) = b_k$ for all k

The algorithm to construct P(x) is called Lagrange Interpolation

Example

$f(x) := 8x^4 + 5x^3 + 12x^2 + 12x + 15;$

$$x \to 8x^4 + 5x^3 + 12x^2 + 12x + 15 \quad (1)$$

$f(0)$ **mod** 29

$$15 \quad (2)$$

$f(1)$ **mod** 29

$$23 \quad (3)$$

$f(2)$ **mod** 29

$$23 \quad (4)$$

$f(3)$ **mod** 29

$$14 \quad (5)$$

$f(4)$ **mod** 29

$$13 \quad (6)$$

$f(5)$ **mod** 29

$$26 \quad (7)$$

$f(6)$ **mod** 29

$$19 \quad (8)$$

CurveFitting[*PolynomialInterpolation*]([[0, 15], [1, 23], [4, 13], [5, 26], [6, 19]], *x*, *form* = *Lagrange*);

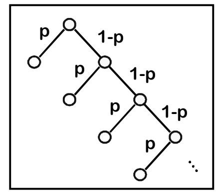
$$\frac{1}{8}(x-1)(x-4)(x-5)(x-6) - \frac{23}{60}x(x-4)(x-5)(x-6) + \frac{13}{24}x(x-1)(x-5)(x-6) - \frac{13}{10}x(x-1)(x-4)(x-6) + \frac{19}{60}x(x-1)(x-4)(x-5) \quad (9)$$

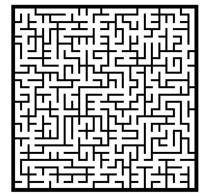
$g :=$ *eval*(*expand*(%) **mod** 29)

$$8x^4 + 5x^3 + 12x^2 + 12x + 15 \quad (10)$$

Infinite Sample spaces and Random Walks

Lecture 17 (October 19, 2010)





Probability Refresher

What's a Random Variable?

A Random Variable is a real-valued function on a sample space S

E[X+Y] = E[X] + E[Y]

Probability Refresher

What does this mean: E[X | A]?

$= \sum_a a \cdot Pr[X = a | A]$

Is this true:

$Pr[A] = Pr[A | B]\, Pr[B] + Pr[A | \overline{B}]\, Pr[\overline{B}]$

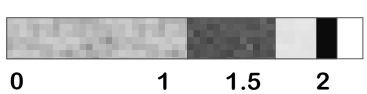
Yes!

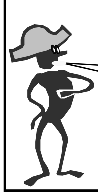
Similarly:

$E[X] = E[X | A]\, Pr[A] + E[X | \overline{A}]\, Pr[\overline{A}]$

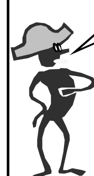
An easy question

What is $\sum_{i=0}^{\infty} (½)^i$? Answer: 2





But it never actually gets to 2. Is that a problem?



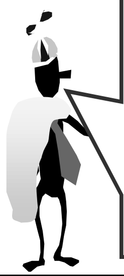
But it never actually gets to 2. Is that a problem?

No, by $\sum_{i=0}^{\infty} f(i)$, we really mean $\lim_{n\to\infty} \sum_{i=0}^{n} f(i)$.

if this limit is undefined, so is the sum

In this case, the partial sum is $2-(½)^n$, which goes to 2.

A related question



Suppose I flip a coin of bias p, stopping when I first get heads.

What's the chance that I:

Flip exactly once?
    p
Flip exactly two times?
    (1-p)p
Flip exactly k times?
    $(1-p)^{k-1}p$
Eventually stop?
    1 (assuming p>0)

A related question

Pr(flip once) +
 Pr(flip 2 times) +
  Pr(flip 3 times) +
   ...
   = 1:

$p + (1-p)p + (1-p)^2p + (1-p)^3p + \dots = 1$

Or, using q = 1-p,

$$\sum_{i=0}^{\infty} q^i = \frac{1}{1-q}$$

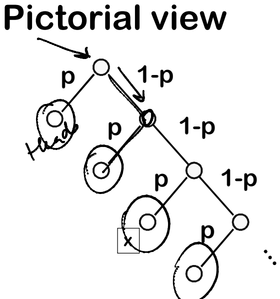
Geometric Random Variable

Flip bias-p coin until you see heads.

Let r.v. Z =
  number of flips until heads

What is E[Z]?

Pictorial view

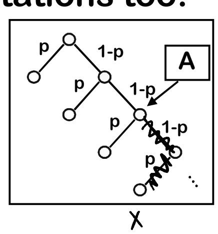


Sample space S = leaves in this tree.
Pr(x) = product of edges on path to x.

$(1-p)^n$

If p>0, Pr(not halting by time n) $\rightarrow$ 0 as $n \rightarrow \infty$.

Reason about expectations too!



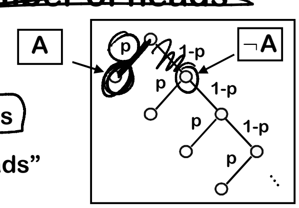
Suppose A is a node in this tree

Pr(x|A)=product of edges on path from A to x.

$E[Z] = \sum_x Pr(x) Z(x)$.

$E[Z|A] = \sum_{x \in A} Pr(x|A) Z(x)$.
 I.e., it is as if we started the game at A.

~~Expected number of heads~~ BAD



Let Z = # flips until heads

A = event "1st flip is heads"

$E[Z] = E[Z|A] \times Pr(A) + E[Z|\neg A] \times Pr(\neg A)$

$= 1 \times p + (1 + E[Z]) \times (1-p)$.

Solving: $p \times E[Z] = p + (1-p)$
 $\Rightarrow E[Z] = 1/p$.

Geom(p) random variable

Z = Number of flips with bias-p coin until you see a heads

E[Z] = 1/p

For unbiased coin (p = ½), expected value = 2 flips

Infinite Probability spaces

Notice we are using infinite probability spaces here, but we really only defined things for finite spaces so far.

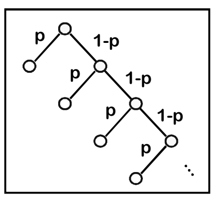
Infinite probability spaces can sometimes be weird.

Luckily, in CS we will almost always be looking at spaces that can be viewed as choice trees where

Pr(haven't halted by time t) $\rightarrow$ 0 as $t \rightarrow \infty$.

A definition for infinite spaces

Let sample space S be leaves of a choice tree.



Let S_n = {leaves at depth $\leq$ n}.

For event A, let $A_n = A \cap S_n$.

If $\lim_{n\to\infty} Pr(S_n)=1$, can define:

$Pr(A)=\lim_{n\to\infty} Pr(A_n)$.

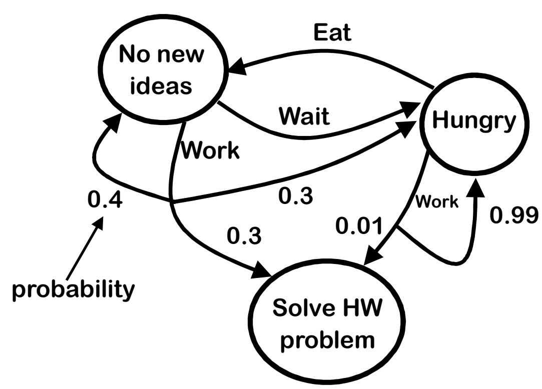
Setting that doesn't fit our model

Event: "Flip coin until #heads > 2 × #tails."

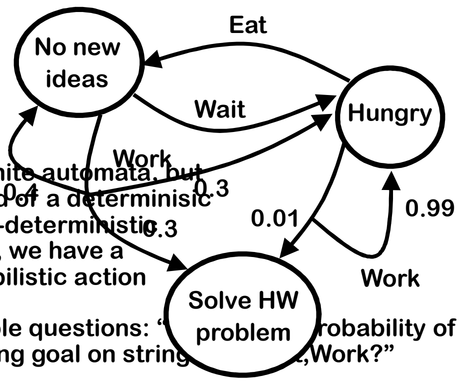
There's a reasonable chance this will never stop...

Random Walks: or, how to walk home drunk

Abstraction of Student Life



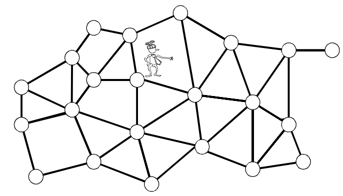
Abstraction of Student Life



Like finite automata, but instead of a determinisic or non-deterministic action, we have a probabilistic action

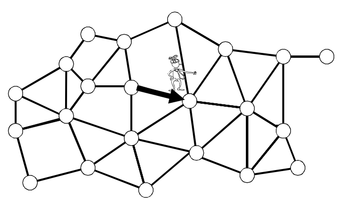
Example questions: “ … robability of reaching goal on string … ,Work?”

Simpler: Random Walks on Graphs



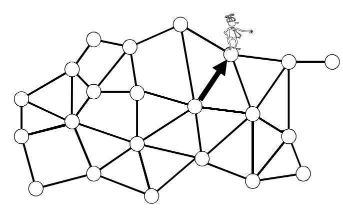
At any node, go to one of the neighbors of the node with equal probability

Simpler: Random Walks on Graphs



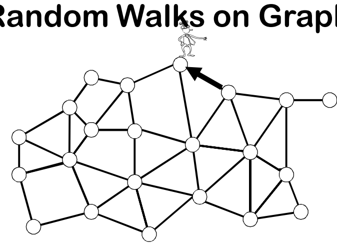
At any node, go to one of the neighbors of the node with equal probability

Simpler: Random Walks on Graphs



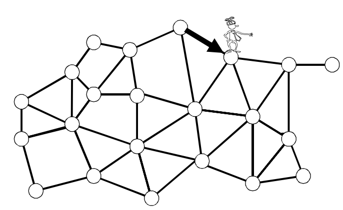
At any node, go to one of the neighbors of the node with equal probability

Simpler: Random Walks on Graphs



At any node, go to one of the neighbors of the node with equal probability

Simpler: Random Walks on Graphs

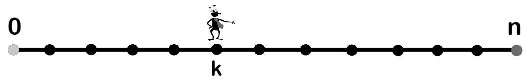


At any node, go to one of the neighbors of the node with equal probability

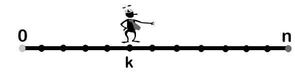
Random Walk on a Line

You go into a casino with \$k, and at each time step, you bet \$1 on a fair game

You leave when you are broke or have \$n



Random Walk on a Line



Question 1: what is your expected amount of money at time t?

Let X_t be a R.V. for the amount of \$\$\$ at time t

Let δ_i be RV for <u>change</u> in money at time i

$E[\delta_i] = 0$ (it's a fair game)

But $X_t = k + \delta_1 + \delta_2 + \ldots + \delta_t,$

So, $E[X_t] = k$

Random Walk on a Line

Question 2: what is the probability that you leave with \$n?

$E[X_t] = k$

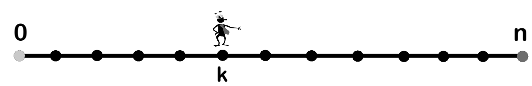
$E[X_t] = E[X_t | X_t = 0] \times \Pr(X_t = 0)$
$\quad + E[X_t | X_t = n] \times \Pr(X_t = n)$
$\quad + E[X_t | \text{neither}] \times \Pr(\text{neither})$

$k = n \times \Pr(X_t = n)$
$\quad + (\text{something}_t) \times \Pr(\text{neither})$

As $t \to \infty$, $\Pr(\text{neither}) \to 0$, also $\text{something}_t < n$
Hence $\Pr(X_t = n) \to k/n$

Another way to see it

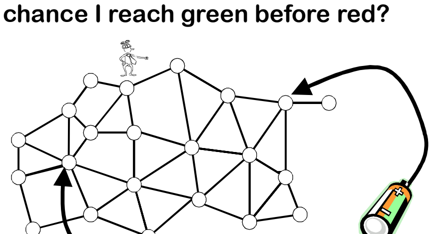
Question 2: what is the probability that you leave with \$n?



= probability that I hit green before I hit red

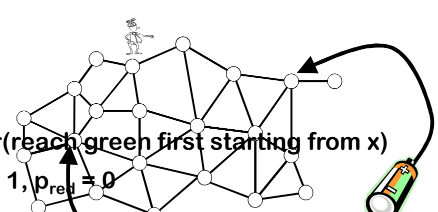
Random Walks and Electrical Networks

What is chance I reach green before red?



Same as voltage if edges are resistors and we put 1-volt battery between green and red

Random Walks and Electrical Networks



p_x = Pr(reach green first starting from x)

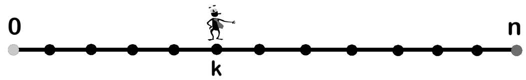
$p_{green} = 1, p_{red} = 0$

And for the rest $p_x = \text{Average}_{y \in Nbr(x)}(p_y)$

Same as equations for voltage if edges all have same resistance!

Another way to see it

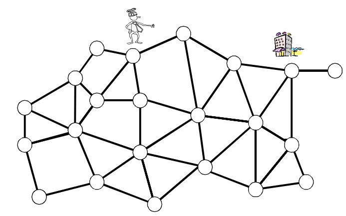
Question 2: what is the probability that you leave with \$n?



= probability that I hit green before I hit red

voltage(k) = k/n
= Pr[hitting n before 0 starting at k] !!!

Getting Back Home

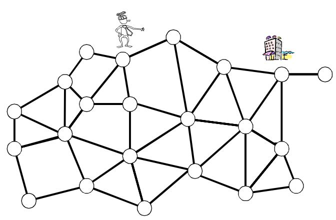


Lost in a city, you want to get back to your hotel
How should you do this?

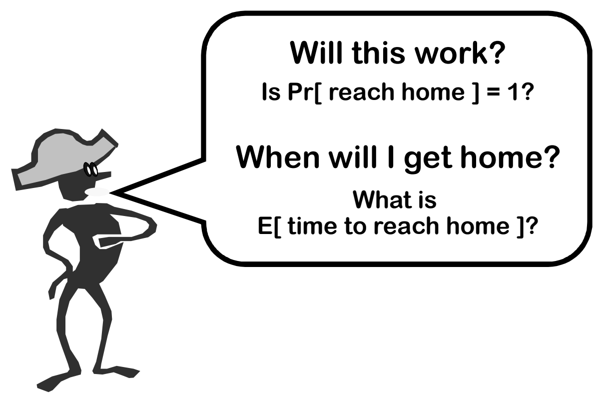
Depth First Search!

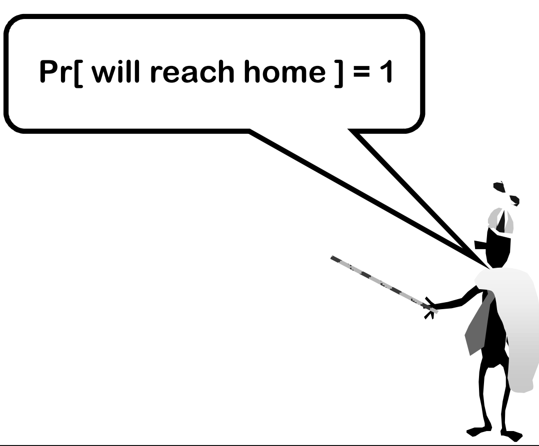
Requires a good memory and a piece of chalk

Getting Back Home



How about walking randomly?





We Will Eventually Get Home

Look at the first n steps

There is a non-zero chance p_1 that we get home

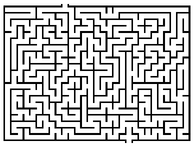
In fact, $p_1 \geq (1/n)^n$

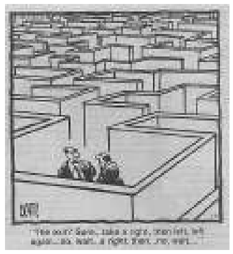
Suppose we don't reach home in first n steps

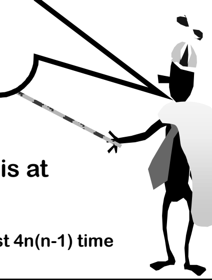
Then, wherever we are, there is a chance $p_2 \geq (1/n)^n$ that we hit home in the next n steps from there

Probability of failing to reach home by time kn
$= (1 - p_1)(1 - p_2) \ldots (1 - p_k) \to 0$ as $k \to \infty$

Getting out of mazes







Theorem:

If the graph has n nodes and m edges, then

E[time to visit all nodes] $\le$ 2m × (n-1)

We will not prove this theorem today

E[time to reach home] is at most this

In a 2-d maze with n intersections, at most 4n(n-1) time

Actually, we get home pretty fast…

Chance that we don't hit home by (2k)2m(n-1) steps is $(\frac{1}{2})^k$

Even if we know the fact on the previous slide, how does one prove this?

A Simple Calculation

True or False:

If the average income of people is $100 then more than 50% of the people can be earning more than $200 each

False! else the average would be higher!!!

Markov's Inequality

If X is a non-negative r.v. with mean E[X], then

Pr[X > 2 E[X]] $\le$ ½

Pr[X > k E[X]] $\le$ 1/k

Andrei A. Markov

Markov's Inequality

Non-neg random variable X has expectation μ = E[X]

μ = E[X] = E[X | X > 2μ] Pr[X > 2μ] + E[X | X $\le$ 2μ] Pr[X $\le$ 2μ]

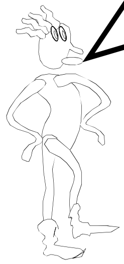
$\ge$ E[X | X > 2μ] Pr[X > 2μ]   (since X is non-neg)

Also, E[X | X > 2μ] > 2μ

$\Rightarrow$ μ $\ge$ 2μ × Pr[X > 2μ]

$\Rightarrow$ ½ $\ge$ Pr[X > 2μ]

Pr[X > k × expectation] $\le$ 1/k



Actually, we get home pretty fast…

Chance that we don't hit home by (2k)2m(n-1) steps is $(½)^k$

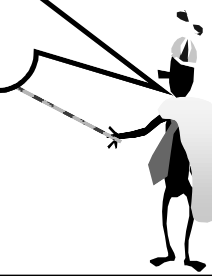
Let's prove this now…

Recall:

If the graph has n nodes and m edges, then

E[time to visit all nodes] $\leq$ 2m × (n-1)

call this value T



An Averaging Argument

Suppose I start at u

E[time to hit all vertices | start at u] $\leq$ T

Hence, by Markov's Inequality:
Pr[time to hit all vertices > 2T | start at u] $\leq$ ½

So Let's Walk Some Mo!

Pr [time to hit all vertices > 2T | start at u] $\leq$ ½

Suppose at time 2T, I'm at some node with more nodes still to visit

Pr [haven't hit all vertices in 2T more time | start at v] $\leq$ ½

Chance that you failed both times $\leq$ ¼ = $(½)^2$

Hence,
Pr[havent hit everyone in time k × 2T] $\leq$ $(½)^k$

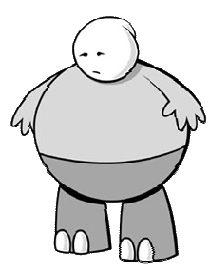
Hence, if we know that

Expected Cover Time

C(G) < 2m(n-1)

then

Pr[home by time 4k m(n-1)] $\geq$ 1 – $(½)^k$



Here's What You Need to Know…

- Conditional expectation
- Flipping coins with bias p
- Expected number of flips before a heads
- Random Walk on a Line
- Cover Time of a Graph
- Markov's Inequality